ORCL-0010801 US P

UNITED STATES PATENT APPLICATION FOR

SIX SIGMA ENABLED WEB-BASED BUSINESS INTELLIGENCE SYSTEM

Inventors:

Timothy M. Dubois
Jacques Senchet
Craig Martell and
Ned Nedumaran

Prepared by:

WAGNER, MURABITO & HAO LLP
TWO NORTH MARKET STREET
THIRD FLOOR
SAN JOSE, CALIFORNIA 95113
(408) 938-9060

SIX SIGMA ENABLED BUSINESS INTELLIGENCE SYSTEM

RELATED U.S. APPLICATION

This Application claims priority to U.S. Provisional Application entitled, "Six Sigma Enabled Business Intelligence System," Application Number 60/248,267,

5 filed on November 13, 2000, which application is hereby incorporated by reference.

FIELD OF THE INVENTION

The present invention relates to the field of business performance

10 measuring. Specifically, the present invention relates to a system for providing sophisticated statistical analysis via a web-based business intelligence system.

BACKGROUND ART

To survive today companies must continuously improve quality and reduce

15 the variability in every process throughout the organization. Access to data and the capability to analyze that data are the keys to reducing defects and increasing customer satisfaction, thus leading to increased productivity and greater top and bottom line growth.

20 Conventional business intelligence systems (BIS) allow a company to view and analyze performance measures across a business enterprise. These systems have useful functions such as: an ability to view data in a graphical or tabular format, to drill down to lower levels of data, and to alert the user if a performance measurement goes outside a target value. Some conventional systems are web-

25 based and provide a configurable home page. Some conventional systems may

have a library of performance measures or key performance indicators to benchmark actual performance against multiple targets. However, these conventional systems do not have the capability to present complex statistical analysis in a user friendly format, which is crucial to reducing defects and

5 improving processes.

Referring to Figure 1, a typical BIS report offers a view of a given performance measure, for example, book-to-ship cycle time. It provides a standard bar or line graph 170 to view the actual data 150 and user defined

10 targets 152. The variance, however, is only listed in the table 160 below the graph 170, which also includes sums or averages of each column. To allow the user flexibility, there are many dimensions 154 (e.g., time, geography, etc.) by which to filter and view the report.

15 Viewing performance measures with line graphs, bar charts, pie charts, etc. provides a standardized way of analyzing data, but in many scenarios is not enough to get a complete and thorough understanding of the actual process capability. Often, more advanced statistics are required to understand the entire issue or discover opportunities that would normally be overlooked. For example,

20 because the average is within the target range, the data may appear acceptable. However, many data points may actually be outside of the target range. Thus, problem data may be overlooked.

There also exist conventional ad-hoc statistical software packages;

25 however, these packages have limitations as well. Such statistical software

ORCL-0010801 US P

packages may be able to assist in measuring the quality of a given process or product. For example, they may calculate a sigma value, which is a statistical parameter that corresponds to a standard deviation on a bell curve. The number of deviations between the statistical mean and the customer defined limits of

5   acceptability provides a quantifiable measurement of process performance. As the process capability increases, so does the sigma level of the process. A process that performs at the six sigma level has six standard deviations between the mean and the customer defined limit, and corresponds to 3.4 defects per million, assuming a normal distribution.

10

However, to perform a six sigma analysis ideally requires involvement of the entire organization, and hence requires access to information relative to a practical measure (e.g., book-to-ship cycle time), which may be distributed across an organization. Unfortunately, these conventional ad-hoc statistical systems do

15   not generally have access to enterprise wide data, without having to download the data to the ad-hoc statistical software package. Downloading such data is often complex, as it may be stored across many databases across the world, and may be in different formats. Thus, the data may also need to be re-formatted before an analysis may be done. Additionally, some conventional statistical packages

20   require the user to have a considerable knowledge of statistics to be able to produce a usable report. Furthermore, the statistical programs themselves may need to be loaded on each computer upon which an analysis is to be run, which is time consuming and costly.

ORCL-0010801 US P

SUMMARY OF THE INVENTION

Therefore, it would be advantageous to provide a method and system for performing a statistical analysis of enterprise wide data over a single pipeline. A further need exists for a method and system which allow the user to be able to

5 have all the functionality of a conventional business intelligence report and seemlessly transition to a detailed statistical analysis of the data under analysis. A still further need exists for a solution which presents a report with a uniform look and feel to all users in a company. An even further need exists for a system which does not require that the user have extensive knowledge of statistics. A still further

10 need exists for a system which pulls data from existing data repositories. An even further need exists for an efficient mechanism for helping a user to understand where to make process improvements by running statistical simulations. A still further need exists for such a system which provides automatic notification to a user when a performance measure goes outside of a user specified target range.

15

The present invention provides a method and system for performing a statistical analysis of enterprise wide data over a single pipeline and allows seamless transitions between a detailed statistical analysis of the data under analysis and more conventional BIS analysis. The present invention presents a

20 report with a uniform look and feel to all users in a company and does not require that the user have extensive knowledge of statistics. The present invention efficiently re-uses data from existing data repositories. The present invention provides for statistical simulations and automatic notifications when a performance measure goes outside of a user specified target range.

A method of presenting an analysis of enterprise wide business data is disclosed. In response to a user request to a web site operable to access enterprise wide business data and to provide statistical analysis, a six sigma

5   enabled BIS transfers an electronic document to the user. The electronic document allows the user to select dimensions to specify which data to analyze for a given performance measure. Then, in response to a request from the user for a statistical analysis, the six sigma enabled BIS performs a statistical analysis of the performance measure and transfers an electronic copy of the statistical analysis to

10   the user. In one embodiment the document comprises a histogram. The histogram has an overlay indicator of the statistical mean, as well as a target limit. In this fashion, the user may view the variance in the data, and see how much data are outside the target range.

15   Other embodiments allow the user to run simulations to see the effect changing a statistical parameter such as the mean or standard deviation will have on the sigma value. In this fashion, the user may see how to best focus improvement efforts. Other embodiments provide a system which provides, in addition to the graphical statistical analysis, an automatic notification when a

20   performance measure goes outside of a user selected target.

The present invention provides these advantages and others not specifically mentioned above but described in the sections to follow.

ORCL-0010801 US P

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is an illustration of a bar graph display of business data, according to prior art.

5   Figure 2 is an illustration of a diagram illustrating data flow, according to embodiments of the present invention.

Figure 3 is a diagram of a histogram comprising statistical analysis of enterprise wide business data, according to embodiments of the present invention.

10

Figure 4 is a diagram of a simulation used to assist process improvement, according to embodiments of the present invention.

Figure 5 is a graph of time versus sigma value, according to embodiments

15   of the present invention.

Figure 6 is a flowchart of the steps of process of performing a statistical analysis of enterprise-wide business data, according to embodiments of the present invention.

20

Figure 7 is an exemplary web-page comprising a histogram of statistical data, according to embodiments of the present invention.

Figure 8 is an exemplary alert notification, according to embodiments of the

25   present invention.

ORCL-0010801 US P

Figure 9 is a schematic of a computer system, which may be used to implement embodiments of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

In the following detailed description of the present invention, a method and system for a six sigma enabled business intelligence report, numerous specific details are set forth in order to provide a thorough understanding of the present

5   invention. However, it will be recognized by one skilled in the art that the present invention may be practiced without these specific details or with equivalents thereof. In other instances, well known methods, procedures, components, and circuits have not been described in detail as not to unnecessarily obscure aspects of the present invention.

10

NOTATION AND NOMENCLATURE

Some portions of the detailed descriptions which follow are presented in terms of procedures, steps, logic blocks, processing, and other symbolic representations of operations on data bits that can be performed on computer

15   memory. These descriptions and representations are the means used by those skilled in the data processing arts to most effectively convey the substance of their work to others skilled in the art. A procedure, computer executed step, logic block, process, etc., is here, and generally, conceived to be a self-consistent sequence of steps or instructions leading to a desired result. The steps are those requiring

20   physical manipulations of physical quantities. Usually, though not necessarily, these quantities take the form of electrical or magnetic signals capable of being stored, transferred, combined, compared, and otherwise manipulated in a computer system. It has proven convenient at times, principally for reasons of common usage, to refer to these signals as bits, values, elements, symbols,

25   characters, terms, numbers, or the like.

ORCL-0010801 US P

It should be borne in mind, however, that all of these and similar terms are to be associated with the appropriate physical quantities and are merely convenient labels applied to these quantities. Unless specifically stated otherwise

5 as apparent from the following discussions, it is appreciated that throughout the present invention, discussions utilizing terms such as "indexing" or "processing" or "computing" or "translating" or "calculating" or "determining" or "scrolling" or "displaying" or "recognizing" or "generating" or the like, refer to the action and processes of a computer system, or similar electronic computing device, that

10 manipulates and transforms data represented as physical (electronic) quantities within the computer system's registers and memories into other data similarly represented as physical quantities within the computer system memories or registers or other such information storage, transmission or display devices.

15 SIX SIGMA ENABLED BUSINESS INTELLIGENCE SYSTEM

The present invention provides the many advanced statistical measures of six sigma analysis in a business intelligence report. A business intelligence system (BIS) may allow, via any web-browser, the ability to analyze performance measures. Performance measures are all the sets of measures that a company

20 would use to access the performance of their enterprise, architectures, or processes. The present invention provides a single pipeline for a multidimensional view of data with a thorough understanding of variation and process capability. In this fashion, key business measures may be comprehensively analyzed in real-time for enterprise wide data.

Definitions

The following definitions may be applied throughout this application.

Actual Value: The calculated value for a performance measure.

5   Collection programs populate summary tables with calculated values and/or the values will be calculated in views. Alerts compare the actual value to the targets in the Performance Management Framework. This actual value will display on the Personal Homepage and on Notification messages.

10   Dimension: A Dimension is a logical entity with respect to which data may be organized and analyzed. For example, data may be organized by dimensions such as 'Time', 'Item', or 'Trading Partner'.

Dimension Level: The level of granularity for a dimension used in
15   selecting and sorting data within the BIS reports, workbooks and performance management framework. Dimension levels form a hierarchy of specificity that together identify the dimension. For example, with the Time dimension, Year, Quarter, and Month are the different dimension levels.

20   Performance Management Framework: The BIS component that allows implementation of performance measure tracking. The framework is used to identify and store the performance measure targets and tolerances, alerts to calculate actual values, and workflow to take action on "out of tolerance" situations.

Performance Measure: An indicator used to determine enterprise performance and success. A company will identify relevant performance measures based on the dynamics of the industry and the company's strategic

5 and tactical goals. This term is often used in conjunction with Balanced Scorecard efforts. This may also be known as a Key Performance Indicator (KPI). The performance measures can be viewed by their standard unit of measure or by a sigma value.

10 Target: A user specified value for a specific performance measure at a certain dimension level and a specific dimension. For example, the target value of 12 might be set for the performance measure Inventory Turns, for the target level of Internal Organization and Time (Financial Period), and the specific dimensions of U.S. Manufacturing and calendar month. For a six sigma

15 enabled business intelligence related report, the Upper and Lower Target Limits may also be defined for the sigma calculation.

Target Level: Used with the Performance Management Framework to identify the dimension level associated with a specific target setup. For

20 example, in setting up a target for the Inventory Turns performance measure, the target level might be Internal Organization and Time (Financial Period).

Tolerance and Lower Target Limit (LTL) and Upper Target Limit (UTL): Used in the Performance Management Framework with Targets to specify a range

25 or boundary around the Target value itself. For example, if the Target for Inventory

ORCL-0010801 US P

Turns is 10, the tolerance might be set at +/- 10% around 10. In this case, the tolerance would be from 9 to 11.

Figure 2 shows the flow of data in a system 200 in which the present

5   invention may be practiced. The six sigma enabled business intelligence system 204 has access to a data warehouse 206. The transactional databases 210 store information on various business transactions, such as orders, shipping dates, revenue, inventory, etc. The transactional databases 210 may reside in different countries and may format the data differently. Commercially available BIS

10   software from Oracle Corporation, Redwood Shores, CA is able to collect the data from the various databases 210 and store them in the data warehouse 206 under common dimensions, a common currency, and a common calendar, etc.

Additionally, the commercially available BIS software is able to run

15   reports against the data in the warehouse 206. The commercially available software is able to identify the data set to be analyzed, the dimensions, and the targets. The BIS may then, for example, compare actual shipping dates against targeted shipping dates.

20   A user is able to access the six sigma enabled BIS 204 via the Internet 54, in the preferred embodiment. Any number of users may access the BIS 204 through any method to access the Internet 54 (e.g., personal computer 202). In the preferred embodiment, the user runs a web-browser to interface with the six sigma enabled BIS 204. However, the access to the six sigma enabled BIS 204

25   is not limited to Internet access. The present invention provides a standard

output for all users, and everyone in the company may see the same basic printout. Therefore, a six sigma analysis is seemlessly deployed across the entire enterprise. Furthermore, the present invention provides a one-touch solution. For example, after a user selects the various dimensions 154 to view

5   in a graph 170, the user may seemlessly transition to a six sigma analysis with a single click.

Six sigma enabled reports portray data in a histogram (e.g., bell curve) format, which displays the variation of the data along with various other

10   statistical parameters. Analyzing data with six sigma enabled BIS 204 provides more details of the data and provides a better understanding of what is driving the higher level "averages" and variation.

Figure 3 illustrates an example of a six sigma enabled BIS report regarding

15   book-to-ship cycle time. The book-to-ship cycle time may be viewed by time, customer, etc. For example, the dimension 154 'year' is 'fiscal year 1998 - fiscal year 1999, and the 'item' is 'laptop', etc. The user may also select the upper and lower target limits, the number of bins (e.g., bars), and the range. Furthermore, the 'view by' 306 selection is set to statistics. When the user hits the 'run' button 308,

20   the statistical analysis is presented as the histogram 302.

The calculations are higher level summaries of the book-to-ship cycle time for each sales order line. In this example, the book-to-ship cycle time for Company B is an average of 25 days. Because the upper target limit (UTL) of 26 days is

25   greater than the average of 25, it may appear that the process is under control.

ORCL-0010801 US P

However, closer analysis using the six sigma enabled BIS 204 of the present
invention, reveals that the distribution (or variation) of book-to-ship cycle times
range from 20 to 30 days and that 27% of the cycle times are above the UTL,
which could potentially lead to customer dissatisfaction. The low sigma value (2.1)
5   also indicates that considerable data lie above the target limit (UTL). The percent
of data out of range and the sigma value may be based on an assumption of a
normal data distribution, although the present invention is not limited to this
assumption.

10   The bars or bins in the histogram 302 which lie outside the acceptable
target limits may be highlighted. For example, in Figure 3, those to the left of the
target limit (UTL) may be displayed in yellow, while those to the right may be
displayed in red to indicate a problem. Thus, the user sees the problem data
which a bar graph 170 such as in Figure 1 fails to reveal.

15

Figure 3 also shows a statistical data table 310, which may be displayed
below the histogram 302. This table 310 shows statistical parameters such as: the
statistical mean, standard deviation, median, actual percentage of data above and
below a user specified target, the 95 percent confidence interval, the number of
20   data points (n), the user specified targets, the Anderson-Darling value, the number
of defects per one million that are out of range (defects per million opportunities or
DPMO), and the sigma value.

Referring now to Figure 4, an embodiment allows the user to run
25   simulations on the data from the report (e.g., the data in Figure 3). For example,

the user may input a hypothetical statistical mean, standard deviation, upper target limit, or lower target limit and run a simulation to produce a new histogram. This embodiment quickly recalculates the data via a simulation of a normal distribution. Thus, this provides a quick way of modeling the changes in the target limits, mean,

5   or variation (standard deviation) and understanding the corresponding impact on the sigma value, confidence interval, risk factors, etc. In this fashion, the user gains insight as to where to focus improvement efforts. For example, it may be more beneficial to attempt to change the mean (average) than to reduce the variation of a process of the actual data.

10

The simulation in Figure 4 determines the effect of reducing the average book-to-ship cycle time to 21 days, while holding the standard deviation at 1.7. In this case, the sigma value would improve from 2.1 to 4.5. Furthermore, the number of defects per million would decrease from 273,451 to 1,337. In a similar

15   fashion, the user may hold the mean constant while testing a hypothetical standard deviation. The present invention is well suited to simulation with varying other statistical parameters.

Referring now to Figure 5, an embodiment produces a graph 500 of the

20   sigma value versus a selected dimension (e.g., time, geography, product, etc.) In this fashion, the user may quickly see the trend of sigma over time.

Figure 6 shows a process 600 of providing a six sigma enabled business intelligence system (BIS) 204. In step 610, the web-site for the six sigma enabled

25   BIS 204 receives a user login request. The six sigma enabled BIS 204 is able to

access data from across an enterprise. For example, the six sigma enabled BIS _ 204 has access to a data warehouse 206 that contains enterprise wide data from databases 210.

5   In step 620, the web-site for the six sigma enabled BIS 204 transfers a Hyper-Text Markup Language (HTML) document such as the exemplary personal homepage 700 shown in Figure 7. As is apparent from Figure 7, the six sigma enabled BIS 204 provides more functionality than only six sigma analysis. For example, links exists in the favorites 702 and navigate 704 boxes on the left of the

10   homepage 700. The web-page 700 may have other formats.

In step 630, the user selects the data set to be analyzed. For example, Figure 3 comprises selectable dimensions 154. The set of books selected is 'Vision Operations (USA)', the trading partner is company B, the item is 'laptop',

15   etc. The user may also select the target limits, number of bins and range. However, default values are provided if these fields are left blank.

In step 640, the user selects the type of view 306 (e.g., statistical, bar graph, etc.). The user may optionally, switch between views 306. In this fashion, the user

20   is allowed a seamless transition between a six sigma data presentation (e.g., histogram 302) and a conventional business intelligence system presentation. The six sigma enabled BIS 204 may use the same database (e.g., data warehouse 206) to create all data presentations. Thus, the present invention makes efficient use of resources.

In step 645, the user selects the performance measure to be analyzed. Performance measure may be defined as an indicator used to determine enterprise performance and success, for example, book-to-ship cycle time.

5

In step 650, in response to the user request for a statistical analysis, the six sigma enabled BIS 204 accesses the enterprise data warehouse 206 and performs a statistical analysis of the user selected parameters.

10   In step 660, the six sigma enabled BIS 204 transfers an HTML document to the user. The document may be one such as shown in Figure 3. In this fashion, the user receives a six sigma analysis from within a BIS, such that the user may seemlessly transition between the various capabilities that the BIS provides, in addition to six sigma analysis.

15

In optional step 670, the user requests that a simulation be run on the data. In response to the request for a simulation, the six sigma enabled BIS 204 performs the simulation and transfers the results to the user, in step 680.

20   Referring again to Figure 7, the HTML page contains a performance measure section 706, including the sigma value for the data in the histogram 302. The HTML page also contains a notifications section 708, which contains links to open a notification. An exemplary notification 800 is seen in Figure 8.

ORCL-0010801 US P

Figure 9 illustrates circuitry of computer system 100, which may form a

platform for embodiments of the present invention. Computer system 100 includes

an address/data bus 99 for communicating information, a central processor 101

coupled with the bus for processing information and instructions, a volatile

5   memory 102 (e.g., random access memory RAM) coupled with the bus 99 for

storing information and instructions for the central processor 101 and a non-

volatile memory 103 (e.g., read only memory ROM) coupled with the bus 99 for

storing static information and instructions for the processor 101. Computer system

100 also includes an optional data storage device 104 (e.g., a magnetic or optical

10   disk and disk drive) coupled with the bus 99 for storing information and

instructions.

With reference still to Figure 9, system 100 of the present invention also

includes an optional alphanumeric input device 106 including alphanumeric

15   and function keys is coupled to bus 99 for communicating information and

command selections to central processor unit 101. System 100 also optionally

includes a cursor control device 107 coupled to bus 99 for communicating user

input information and command selections to central processor unit 101.

System 100 of the present embodiment also includes an optional display

20   device 105 coupled to bus 99 for displaying information. A signal input/output

communication device 108 coupled to bus 99 provides communication with

external devices.

Exemplary Calculations for Statistical Parameters

Following are calculations for statistical parameters, although the present invention is well-suited to using other formulas.

5   Mean: The Mean is the average of the data set. This calculation will be based off of the data that is defined by the user's selection of dimensions and parameters. Mean = Sum of data points/number of data points.

Standard Deviation: The Standard Deviation is the measure of variation
10   for the given data set. The formula is given by: SQROOT$(((n*Sum(x^2))$ - $(Sum(x))^2)/n*(n-1))$

Median: The Median represents the middle of the data set. For example, Median (1, 2, 3, 4, 5) = 3 and the Median (1, 2, 3, 5) = 2.5

.15

Lower Target Limit (LTL) and Upper Target Limit (UTL): The Lower Target Limit and Upper Target Limit are not calculations. They are user defined in two possible ways.

1. Define Target Limits in PMF (Performance Management Framework).
20   The upper and lower ends of the range from PMF represent the UTL and the LTL. The Upper and Lower Target Limits will be defined in two possible ways:

a. Input a Tolerance % around the Target. This will automatically fill in   the Upper and Lower Target Limit numbers.

b. Input the Upper and Lower Target Limit numbers. This will
25   automatically fill in the Tolerance %.

2. Define the UTL and LTL for the simulation. The simulation area lets the user override the UTL and LTL brought in from PMF or enter them as new values if Targets in PMF were not set up.

5

Actual < LTL and Actual > UTL: The Actual < LTL is the Actual % of data that is below the Lower Target Limit. The Actual > UTL is the Actual % of data that is above the Upper Target Limit: Actual < LTL= (# of data points below the LTL)/Total # of data points * 100

10

95% CI (95% Confidence Interval): The 95% Confidence Interval of the data set. It is the prediction interval where 95% of the data will fall between. 95% CI = Mean +/- 1.96 * Standard Deviation. For example, 95%CI where (Mean = 5, Stdev = 1) => 5 +/- 1.96 * 1 => 3.04 to 6.96

15

Range of Data: The Range of Data is simply the lowest value to the highest value in the given data set. The actual lowest and highest values will be returned for these fields, but may be overrided by inputting values. The reason is to eliminate outliers and recalculate.

20

# of Bins: The # of Bins (also known as Buckets or Classes) refers to the # of Bins in the Histogram (the bars that make up the graph). This is somewhat of a controversial issue since having too few Bins could cover up a trend and too many could dilute any trends. The rule of thumb is to choose between 5 and 20 Bins for any given data set. Following is a simple matrix to use:

25

ORCL-0010801 US P

# of different data points	# of bins
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12-100	11
101-200	15
> 200	20

n (Sample Size): n represents the number of data points in the data set. It is the Sample Size.

A-D Normality (Anderson-Darling Normality Test): The Anderson-Darling Normality test is a test to see if the data is normally distributed. The test will yield a Pvalue that should be larger than .05 for normal data. If the Pvalue is less than .05, it cannot be stated for certain that the data is normal.

$$A^2_n = (-\{\sum_{i=1}^{n} (2i - 1)[\ln Z_i + \ln(1 - Z_{n+1-i})]\}/n - n$$

The Pvalue is found in a corresponding cross reference table.

Probability of being below the LTL: The probability of being below the LTL is based off of a normal distribution and the Z deviate (Z table). To calculate the Probability, the Z deviate is first calculated as:

$Z = (X - LTL)/Stdev$ [(Mean - Lower Target Limit) divided by the Standard

5   Deviation

Once the Z deviate is found, the probability can be found by looking it up in the Z table which is found in any statistics book.

Probability of being above the UTL: The probability of being above the

10   UTL is based off of a normal distribution and the Z deviate (Z table). To calculate the Probability, the Z deviate is first calculated as:

$Z = (UTL - X)/Stdev$ [(Upper Target Limit - Mean) divided by the Standard Deviation

Once the Z deviate is found, the probability can be found by looking it up in the

15   Z table which is found in any statistics book.

Sigma or Z (Short Term): The value is calculated from the Z table. The "Probability of being above the UTL" is added to the "Probability of being below the LTL". This Total Probability is looked up on the Z table (area under the curve)

20   and related back to a Z deviate. This value represents the Long Term Sigma. A common rule of six sigma Methodology says that a Short Term Sigma can be represented as the Long Term Sigma + a 1.5 sigma shift. This reasoning is based on the fact that over a long period of time, the distribution will spread out due to time related sources of error which tend to upset process centering as compared

25   to the Short Term Data gathering/analysis. The "Six Sigma Level of Excellence" is

ORCL-0010801 US P

based off of the short term (Zst). If your data is short term in nature, then the resultant sigma value will be short term. If your data is long term in nature, the Zst is calculated off of the Long Term Sigma + 1.5. So, Zst = Zlt + 1.5 shift.

5   Guidelines for the Correction of Sigma Values (Excerpt from "The Vision of Six Sigma: Tools and Methods for Breakthrough" by Mikel J. Harry - 1994

       Guideline 1: If a metric is computed on the basis of data gathered over many cycles or time intervals, the resultant value should be regarded as a long-
10   term measure of performance. Naturally, the long-term metric must be converted to a probability. Once expressed as a probability, Z.lt value may be established by way of a table of area-under-the-normal-curve, or any acceptable computational device. If we seek to forecast short-term performance (Z.st), we must add a shift factor (Z.shift) to Z.lt so as to remove time related sources of
15   error which tend to upset process centering. Recognize that the actual value of Z.shift is seldom known in practice; therefore, it may be necessary to apply the accepted convention and set Z.shift = 1.50; otherwise, use the actual value. As a consequence of this linear transformation, the final Z value should reflect only random sources of error and, therefore, serve as a projection of short-term
20   performance. Thus, we are able to artificially remove the effect of non-random influences (i.e., normal process centering errors) from the analysis via the transform Z.st = Z.lt + Z.shift.

       Guideline 2: If the metric is computed on the basis of data gathered over
25   a very limited number of cycles or time intervals, the resultant value should be

regarded as a short-term measure of performance. Naturally, the short-term metric must be converted to a probability. Once expressed as a probability, Z.st may be established by way of a table of area-under-the-normal-curve, or any acceptable computational device. If we seek to forecast long-term performance,

5   we must subtract Z.shift from the Z.st so as to approximate the long-term capability. Recognize that the actual value of Z.shift is seldom known in practice; therefore, it may be necessary to apply the accepted convention and set Z.shift = 1.50. If the actual value is known, use it. As a consequence of this linear transformation, the final Z value reflects both random and non-random sources

10   of error and, therefore, is a projection of long-term performance. Thus, we are able to artificially induce the effect of non-random influences (i.e., normal process centering error) into the analysis by way of Z.st - Z.shift = Z.lt.

Guideline 3: In general, if the originating data is discrete by nature, the

15   resulting Z transform should be regarded as long-term. The logic of this guideline is simple; a fairly large number of cycles or time intervals is often required to generate enough non-conformities from which to generate a relatively stable estimate of Z. Hence, it is reasonable to conclude that both random and non-random influences are reflected in such data. In this instance,

20   guideline 1 would be applied.

Guideline 4: In general, if the originating data is continuous by nature and was gathered under the constraint of sequential or random sampling across a very limited number of cycles or time intervals, the resulting Z value

25   should be regarded as short-term. The logic of this guideline is simple; data

gathered over a very limited number of cycles or time intervals only reflects

random influences (white noise) and, as a consequence, tends to exclude non-

random sources of variation, such as process centering errors.

5   Guideline 5: Whenever it is desirable to report the corresponding "sigma" of

a given performance metric, the short-term Z must be used. For example, let us

suppose that we find 6210 ppm defective. In this instance, we must translate 6210

ppm into its corresponding "sigma" value. Doing so reveals Z.lt = 2.50. Since the

originating data was long-term by nature, guideline 1 and 3 apply. In this case, Z.lt

10   + Z.shift = 2.5 + 1.5 = 4. Since no other estimate of Z.shift was available, the

convention of 1.5 was employed.

       The preferred embodiment of the present invention, a method and system

15   for a six sigma enabled business intelligence system, is thus described. While the

present invention has been described in particular embodiments, it should be

appreciated that the present invention should not be construed as limited by such

embodiments, but rather construed according to the below claims.

ORCL-0010801 US P